Image Transmission over the Polya Channel via Channel-Optimized Quantization

Firouz Behnamfar, *Student Member, IEEE,* Fady Alajaji, *Senior Member, IEEE,* and Tamás Linder, *Senior Member, IEEE*

***Abstract*— We introduce two progressive methods for image transmission over binary channels with additive bursty noise modeled by the finite-memory Polya (contagion) channel. The methods, which are based on channel-optimized scalar quantization (COSQ) of the wavelet transform coefficients, exploit channel memory to offer superior performance over a number of more complex systems designed for the fully interleaved channel.**

***Index Terms*— Subband coding, joint source-channel coding, channel optimized scalar quantization, channels with memory, convolutional codes.**

I. INTRODUCTION

The traditional approach to data compression and channel coding is to treat these two tasks independently. The performance of the resulting systems – often called tandem systems – may be far from optimal when resources (computational/memory complexity and delay) are restricted. As a result, it is beneficial to perform source compression and channel error protection jointly. This issue was quantitatively studied, for example, recently in [15], where joint source-channel coding is shown to outperform tandem coding if the systems have delay or complexity below a certain threshold. Several methods have been proposed for joint source-channel coding of images, including unequal error protection (UEP) [18], [21], [23], [26], [27] and channel-optimized scalar quantization (COSQ) [4]. In [26], image compression is performed through subband coding (with quadrature mirror filter (QMF) banks) followed by variable-length coding, and channel protection is done using packetization, check sum bits, and rate-compatible punctured convolutional (RCPC) codes. The approach in [23] is similar to [26], but the source coder is replaced with the strong method SPIHT [22]. In [21], each subband is divided into a number of blocks. For every block, the bits at each level of significance are encoded with a different RCPC code rate. Finally, in [4], COSQ is used to transmit the transform coefficients. As COSQ is a fixed-length code, it does not suffer from the usual error propagation effect due to the use of lossless variable-length source codes. COSQ also combines source and channel coding, hence it reduces the overall system complexity as compared with UEP methods.

In this correspondence we present two COSQ-based image coders for transmission of images over noisy channels with memory. Memory is an important property of many real life channels and is usually combated using channel interleaving (since most coding techniques are designed for memoryless channels). However, interleaving causes delay and increases complexity. Furthermore, the resulting associated memoryless channel has a lower capacity than the original channel with memory (for the broad class of information stable channels [1]). One salient feature of our COSQ methods is that they take into account the channel intra-block memory [19]; hence their performance improves as channel noise becomes more correlated. Moreover, they provide reasonable image quality at bit rates as low as 0.125 bpp and channel bit error rates (BERs) as high as 0.1. They also outperform a typical UEP scheme which uses scalar quantization and convolutional coding (with and without channel interleaving).

In our systems, the source image is decomposed with a discrete wavelet transform (DWT) or QMF banks into a number of subbands. As we begin the transmission of the data from the lower frequency subbands (and hence the reconstructed image quality improves as more data is received), the systems are progressive. For bit allocation among the subbands, we employ a generalized version of the technique in [28], which has low computational complexity. Finally, in the presence of channel mismatch between the design and actual channel BERs, the employed systems inherit the resilience of COSQ to outperform one of the strongest UEP methods that is introduced in [23]. When there is no channel mismatch, the UEP system of [23] (which employs interleaving) has a better performance than our methods; however, it has larger complexity and delay.

The rest of this correspondence is organized as follows. In Section II, the structure of our image transmission systems is described. In particular, the image decomposition, subband modeling, channel modeling, quantization, and bit allocation steps are addressed in detail. Simulation results are presented in Section III. Finally, conclusions are stated in Section IV.

II. IMAGE CODING FOR THE BINARY POLYA CHANNEL WITH MEMORY

A. Structure

Figure 1 shows the block diagram of the employed image coding systems which use the generic structure of [28]. It is assumed that the mean of the pixel intensities is removed. In the first method, the image is divided into a number of blocks. Every block is then transformed using the 2-dimensional (2-D) DWT three times, every time on the lowest resolution level of the previous decomposition level. If the block size is $L \times L$ (we use 8×8), we next form L^2 sub-sources by grouping all $c_{i,j}^{(k)}$, $k = 1, ..., K$ together, where K is the number of blocks and $c_{i,j}^{(k)}$ is the coefficient at row i and column j of block k. This method will be referred to as InterBlock. In the second method, the image is treated as one block. It is decomposed four times using the 32-D QMF banks of [14] resulting in 13 subbands, each of which is regarded as a sub-source. We shall refer to this system as IntraBlock. In both of these systems, the sub-sources are quantized using a COSQ for channels with memory with rates determined as explained in Subsection II-D. The resulting bit-stream is then sent directly over the channel. The receiver is simply the inverse of the transmitter.

This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada and the Premier's Research Excellence Award (PREA) of Ontario.

Parts of the material in this correspondence were presented at The 2002 IEEE Data Compression Conference (DCC'02), Snowbird, Utah, March 2002, and The 21st Biennial Symposium on Communications, Kingston, Ontario, Canada, June 2002.

F. Behnamfar is with the Department of Electrical and Computer Engineering and F. Alajaji and T. Linder are with the Department of Mathematics and Statistics and the Department of Electrical and Computer Engineering, Queen's University, Kingston, Ontario K7L 3N6, Canada.
E-mail: {firouz, fady, linder}@mast.queensu.ca, Phone: (613) 533-6000 ext. 75069, Fax: (613) 533-6615.

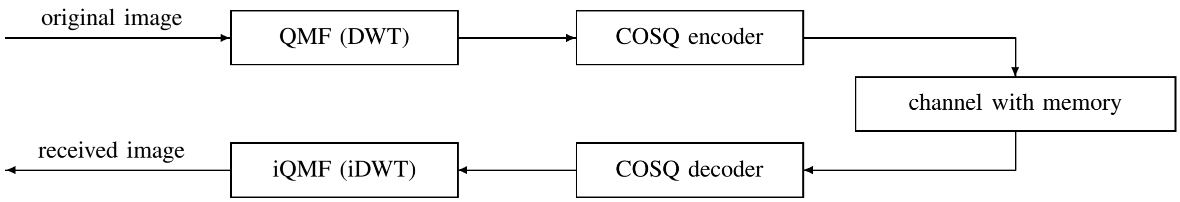

Fig. 1. The structure of the proposed image communication systems.

For COSQ design, we need a good model for the distribution of the samples to be quantized. It is well known that the distribution of the coefficients of every subband is well approximated by the generalized Gaussian distribution [4], [26], with a probability density function given by

$$f(x) = \frac{\alpha\eta(\alpha,\sigma)}{2\Gamma(1/\alpha)}\exp\{-[\eta(\alpha,\sigma)|x|]^{\alpha}\}$$

where $\eta(\alpha,\sigma) = \frac{1}{\sigma}\left(\frac{\Gamma(3/\alpha)}{\Gamma(1/\alpha)}\right)^{\frac{1}{2}}$ is the rate of decay, σ^2 is the variance, and $\Gamma(\cdot)$ is the Gamma function. For $\alpha = 1$ and 2, the above yields the Laplacian and Gaussian distributions, respectively. Based on our experiments on natural images and the results of [4], [26], we assume here that the sub-sources in all subbands have the Laplacian distribution and we quantize them using a COSQ trained for such a source. The benefit of sub-source modeling is that the codebooks will not be image dependent and require no side information.

B. The Polya Contagion Model for Channels with Memory

Most real-world communication channels suffer from correlated noise and/or fading distortions, often occurring in a bursty fashion. The most commonly used binary channel model to represent the equivalent discrete (hard-decision demodulated) fading channel with memory is the well-known Gilbert-Elliott channel (GEC) which was first introduced in [11] and later generalized in [7]. However, this channel model, which is based on a Hidden Markov error process, is often difficult to mathematically analyze as it does not admit a closed-form expression for its channel capacity and its block transition distribution is not transparently expressed in terms of the channel parameters. Recently, a simpler channel model assuming additive channel noise was introduced in [1], where the additive error process is generated by Polya's contagion urn scheme with finite memory. This noise process is a stationary ergodic Markov process of order M. To further motivate the selection of our channel model, we note that it has been recently shown (numerically) that the class of binary channels with additive M^{th} order Markov noise (to which the channel in [1] belongs as a special case) is a good approximation (and, indeed, better than the GEC model) to hard-decision demodulated channels with correlated Rayleigh and Rician fading [20]. The Polya channel with $M = 1$ is identical to an arbitrary channel with stationary ergodic first order Markov noise; hence it can approximate the Rayleigh fading channel in certain conditions [20]. We adopt the finite-memory Polya-contagion channel model in [1] which assumes that any noise sample depends only on the sum of the M previous samples. If X_i, Y_i, and Z_i represent the input, output, and noise in that order and $\oplus$ is addition modulo 2, the channel input-output relationship is described by $Y_i = X_i \oplus Z_i$. Assuming that the input and noise are independent, for $i \geq M$ and any $e_{i-M}^{i-1} \in \{0,1\}^M$, we have, (see [1]):

$$\Pr\{\, Z_i = 1 |\, Z_{i-M}^{i-1} = e_{i-M}^{i-1}\} = \frac{\epsilon + \delta\sum_{j=i-M}^{i-1} e_j}{1 + M\delta}$$

where ϵ is the BER and $\delta \geq 0$ controls the correlation coefficient of the noise given by $\frac{\delta}{\delta+1}$. The channel capacity (whose closed-form expression is derived in [1]) increases with δ, suggesting that channel-optimized vector quantization (COVQ) may achieve less distortion for channels with memory. This intuition was empirically validated by [19], where a COVQ scheme was applied for the compression of ideal sources over this channel model. If δ is set to zero, the noise process becomes memoryless and the channel reduces to a binary symmetric channel (BSC). Note also that this model is less complex than the Gilbert-Elliott channel model [17] and is completely specified with only three parameters (ϵ, δ, and M). It has been recently adopted in several joint source-channel coding studies that include wireless applications (e.g., [5], [12], [13], [25]).

C. COSQ Design

Let $d = d_H(\mathbf{x},\mathbf{y})$ be the Hamming distance between the binary channel input block $\mathbf{x} = (x_1,\ldots,x_n)$ and the output block $\mathbf{y} = (y_1,\ldots,y_n)$. The channel block transition probabilities are given by [1]:

- For $n \leq M$, $\mathrm{P}(\mathbf{y}|\mathbf{x}) = L(n,d,\epsilon,\delta)$, where

$$L(n,d,\epsilon,\delta) = \frac{\prod_{i=0}^{d-1}(\epsilon+i\delta)\prod_{i=0}^{n-d-1}(1-\epsilon+i\delta)}{\prod_{i=0}^{d-1}(1+i\delta)}.$$

- For $n > M$,

$$\begin{aligned}\mathrm{P}(\mathbf{y}|\mathbf{x}) &= L(M, s_{M+1}, \epsilon, \delta)\\ &\quad\times\prod_{i=M+1}^{n}\left[\frac{\epsilon+s_i\delta}{1+M\delta}\right]^{e_i}\left[1-\frac{\epsilon+s_i\delta}{1+M\delta}\right]^{1-e_i}\end{aligned}$$

  where $e_i = x_i \oplus y_i$ and $s_i = e_{i-M} + \cdots + e_{i-1}$.

The significance of the above formulas is that unlike many other channel models in the literature, they provide easy and computationally inexpensive tools to implement the modified generalized Lloyd algorithm (GLA) for noisy channels [9].

Various algorithms have been proposed for COVQ design among which we tried the modified GLA initialized by simulated annealing [8], noisy channel relaxation [10], stochastic relaxation [29], and deterministic annealing [16]. After some experiments with the above algorithms, we decided to use the modified GLA with simulated annealing because of its computational efficiency. Unlike other methods, it computes the codebooks for rates as high as 8 bits per sample in a reasonable time. Also, the training distortion of the quantizers trained with this method were the most consistent with many trials (see Section III for more practical issues).

D. Bit Allocation

In subband coding of images, the end-to-end distortion is more sensitive to errors in the low resolution subbands. Therefore, when allocating bits to the sub-sources, the subbands in which they are located should be taken into account. For the sake of brevity, in what follows we explain the bit allocation strategy for the IntraBlock method only. Bit allocation for the InterBlock method is essentially the same.

Usually, the distortion of sub-source i is weighted by w_i, the L_2 norm of the basis functions of the subband to which it belongs (see, for example, [2] and [4] for the values of w_i when DWT and QMF are used, respectively). Using the mean-square error distortion measure, we can write the end-to-end distortion as

$$D = \sum_{i=1}^{S} w_i d_i, \quad d_i = \frac{1}{N_i^2} \sum_{m=1}^{N_i} \sum_{n=1}^{N_i} (c_{m,n}^{(i)} - \hat{c}_{m,n}^{(i)})^2, \qquad (1)$$

where S is the number of the sub-sources (there are $S = 13$ sub-sources in our implementation of the IntraBlock method), $\hat{c}_{m,n}^{(i)}$ is the reconstructed value of $c_{m,n}^{(i)}$, and we have assumed that the i^{th} subband has a size of $N_i \times N_i$[1]. We employ dynamic programming for bit allocation. In particular, we extend the work in [28] for the Markov noise channel and for the case where the overall distortion has different sensitivities to different sub-sources. The bit allocation problem is to minimize D in (1) subject to $\sum_{i=1}^{13} N_i^2 r_i \leq B$ and $0 \leq r_i \leq r_{\max}$, where r_i is the number of bits per sample allocated to the i^{th} sub-source and B is the total number of bits available. $r_{\max}$ is the maximum number of bits per sample which may be allocated to a sub-source. Note that the overall bit rate is $\frac{1}{RC} \sum_{i=1}^{13} N_i^2 r_i$, where the image size is assumed to be $R \times C$. We choose $r_{\max} = 8$ bits to have small codebooks and fast encoding[2]. Modeling the sub-sources as independent Laplacian sources, we can write each d_i in (1) as $\sigma_i^2 d_L(r_i)$ where $d_L(r_i)$ is the distortion of a unit-variance Laplacian source quantized for a set of channel conditions (i.e., ϵ, δ, M) and σ_i^2 is the variance of the i^{th} sub-source. The problem now is to allocate the available bits to 13 Laplacian sources, each with variance $w_i \sigma_i^2$, given the channel conditions. We use the algorithm in [28] to solve this problem which is guaranteed to achieve the optimal bit allocation.

Note that $d_L(r_i)$ is calculated off-line. Also, although σ_i^2 is image-dependent, it is not computed inside the algorithm. For a 512×512 image, the above algorithm requires a worst case (unlikely) number of 262261 multiplications, 104 additions, and $3072r$ comparisons for the calculation of the σ_i^2 and rate allocation, where r is the overall bit rate. These figures change to 262720 multiplications, 512 additions, and $4032r$ comparisons for the InterBlock method (see the Appendix for the number of operations in general). In both cases, the number of operations is very small compared with those required in UEP methods, where rate allocation forms a significant part of the computations (see, for example, [3], [24])[3].

E. Side Information

Side information is the part of data in which no single bit error can be tolerated. Therefore, this amount must be kept as small as possible[4]. The amount of side information of our IntraBlock method is identical to that in [4]. For a 512×512 image, the IntraBlock method needs the image mean, 13 sub-source variances, and the bit allocation table. Using 10, 10, and 3 bits for each of these amounts to a side information of 0.00068 bits per pixel (bpp). For such an image and block size of 8×8, the InterBlock method has 64 sub-sources and its side information becomes 0.0032 bpp. In comparison to the above figures, the method in [21] requires 0.07 bpp of side information, assuming that 1.5 significance layers are encoded on average and four different RCPC code rates are used.

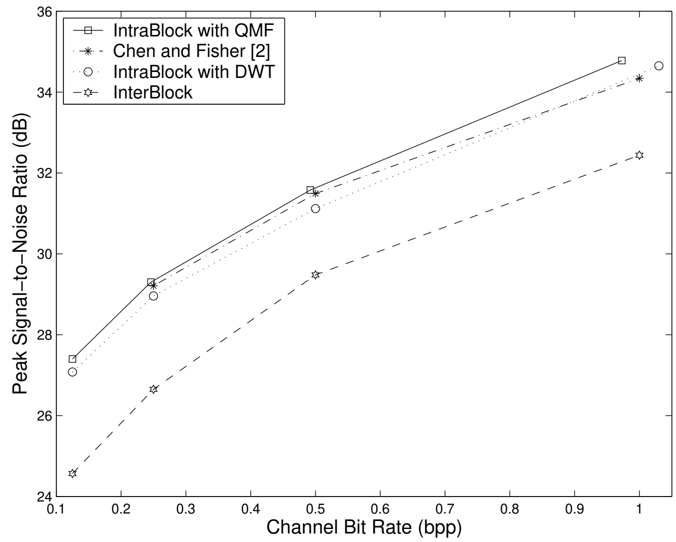

Fig. 2. Performance of various methods at 0.5 bpp, noiseless channel.

III. Simulation Results

We implemented the proposed image coder for the compression and transmission of gray-scale images over the contagion channel with $M = 1$ and tested it for the image Lena (tests performed on other images such as Goldhill, Baboon, and Peppers gave results consistent with the Lena experiments). It takes around twenty hours to train the rate-8 bps COSQ's with 200000 samples on Sun Ultra 60 machines. Over 5 trials, the training distortion was the same within 0.1 dB.

Figure 2 compares "IntraBlock with QMF", which is the IntraBlock method, "Chen and Fisher", which is the method in [4], "IntraBlock with DWT", whose structure is identical to "IntraBlock with QMF" but uses DWT, and InterBlock at an overall bit rate of 0.5 bpp for a noiseless channel. The performance of each system is measured in terms of the peak signal-to-noise ratio (PSNR), defined as PSNR $= 10 \log_{10} \frac{255^2}{\sum_{i,j}(x_{i,j} - \hat{x}_{i,j})^2}$ (dB), where $x_{i,j}$ and $\hat{x}_{i,j}$ are the original and reconstructed pixel values, respectively. For DWT, we used the 9/7 Daubechies filters which are used in the JPEG2000 image compression standard [6]. Figure 2 shows that IntraBlock (with QMF) outperforms the best reported COSQ-based method [4] which has *exactly* the same structure and filters as ours. This is due to our improved bit allocation and selection of the w_i weights in (1). Note that the method in [4] is not suitable for channels with memory and requires channel interleaving. Figure 2 also demonstrates that using QMF banks yields better results than DWT filter banks (at the expense of higher arithmetic complexity).

Figures 3 and 4 compare our methods in noisy channel conditions at a bit rate of 0.25 bpp with a typical UEP system which is denoted by "UEP-IL" and comprises transform coding with QMF banks, scalar quantization of the transform coefficients, convolutional coding, and ideal channel interleaving (the plots for other bit rates show the same behavior). The convolutional codes have 64 states, are non-symmetric, and are of rates $1/2$ ($g_{11} = 634, g_{12} = 564, d_{free} =$

[1]For the InterBlock method, the overall distortion can be written as $D = \sum_{i,j=1}^{L} w_{i,j} d_{i,j}$, where $d_{i,j} = \frac{1}{K} \sum_{k=1}^{K} \left(c_{i,j}^{(k)} - \hat{c}_{i,j}^{(k)} \right)^2$.

[2]Using a larger value for $r_{\max}$ will enhance the performance, but the training time and the size of the required codebooks will grow exponentially and tend to be prohibitive for $r_{\max} > 8$ (given our computational resources).

[3]In order to reduce the complexity of rate allocation for such UEP methods, a parametric approach is developed in [18], where it is assumed that the logarithm of the block error rate has a linear relationship with the reciprocal of the channel code rate.

[4]We assume that the side information can be transmitted without errors using a very strong channel code. Since the amount of the side information is negligible as compared with the image information, this will cause a minimal increase in the overall bit rate.

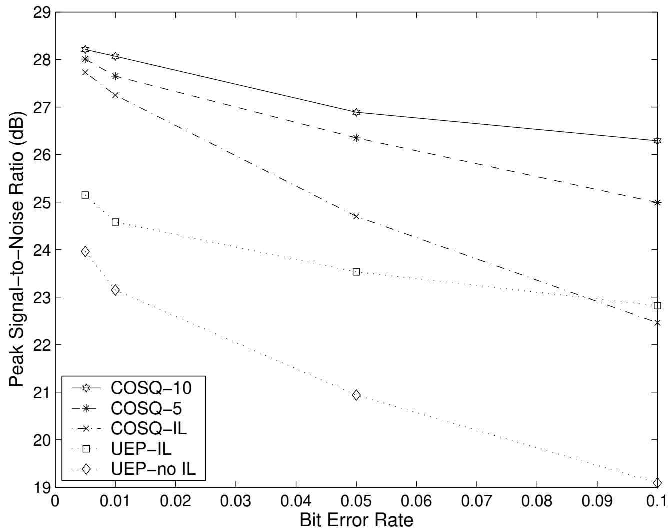

Fig. 3. Typical PSNR performance of various InterBlock methods over the binary channel with memory at 0.25 bpp. Block size is 8×8.

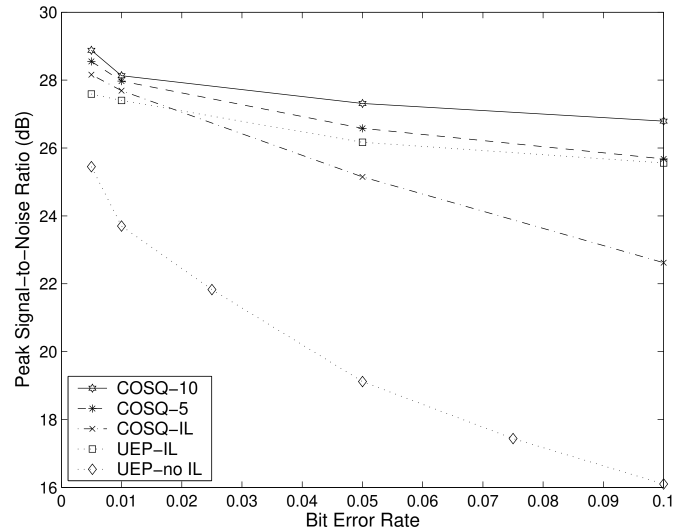

Fig. 4. Typical PSNR performance of various IntraBlock methods over the binary channel with memory at 0.25 bpp.

TABLE I

BER MISMATCH RESULTS IN TERMS OF PSNR IN DB FOR A RATE OF 0.5 BPP AND DESIGN BER OF 0.01.

Actual BER	0	0.01	0.03	0.04	0.05
IntraBlock	31.15	30.43	29.18	28.79	28.25
InterBlock	30.77	30.08	28.81	28.16	27.81
UEP method of [23]	34.97	34.97	28.33	21.90	18.33

10), $1/3$ ($g_{11} = 574, g_{12} = 664, g_{13} = 744, d_{free} = 14$), and $1/4$ ($g_{11} = 454, g_{12} = 574, g_{13} = 664, g_{14} = 724, d_{free} = 20$). The UEP system without the channel interleaver is also implemented and denoted by "UEP-no IL". The UEP systems choose the source and channel coding rates that maximize the PSNR for each set of the channel conditions[5]. We refer to our system as COSQ, followed by the value of the channel correlation parameter (δ) it is designed for (e.g., COSQ-5 and COSQ-10). COSQ-IL denotes the same system which uses an ideal channel interleaver ($\delta = 0$), and hence it is designed for the BSC with the same BER as the channel with memory. The plots show that the performance curves of the image coders designed for the correlated channel are higher than those of the interleaved channel. This demonstrates that substantial gains may be obtained from exploiting the channel memory instead of using interleaving, which increases delay and memory requirements. It is also observed that removing the channel interleaver to reduce the complexity of the UEP system (with the structure used here) would have catastrophic results. The plots also confirm the superiority of the IntraBlock over the InterBlock methods, as was seen in Figure 2 for the noiseless channel.

The output image of three IntraBlock methods for the image Lena at a rate of 0.5 bpp is shown in Figure 5. The figure shows that our method, which exploits channel memory, provides a higher quality decoded image than the systems which use channel interleaving.

So far, our image coding methods assume perfect knowledge of the channel parameters at both transmitter and receiver. However, in real life situations, channel parameters are time-varying, resulting in channel mismatch. Table I compares our two coders in the presence of BER mismatch and that of Sherwood and Zeger [23], which is one of the best UEP schemes reported and is designed for the ideally interleaved channel (in [3] and [27], schemes that outperform [23] have been presented, but these works do not report on mismatch performance). It is observed that the IntraBlock system is the most robust against BER mismatch. Although the substantially more complex Sherwood-Zeger scheme performs better than ours by about 4 dB at no mismatch, it performs nearly 10 dB worse than our method in the presence of mismatch due to the variable-length lossless source code employed by that method. Resilience against mismatch is indeed a property of COSQ-based systems which is inherited by our methods.

[5]The source and channel coding rates were found to be $(1/8, 1/2)$, $(1/8, 1/2)$, $(1/12, 1/3)$, and $(1/16, 1/4)$ for BER = $0.005, 0.01, 0.05$, and 0.1, respectively.

IV. CONCLUSION

Two progressive image coding methods for the Polya contagion channel, which is a binary noisy channel with memory, are presented. Both systems are based on transform coding, subband modeling, and COSQ. Encoding an entire subband is observed to be better than dividing the image into blocks and encoding one coefficient from each block. The main result is that it is more beneficial to exploit the channel memory in the COSQ design as opposed to suppressing its effect through channel interleaving, which also increases delay and complexity. COSQ was shown to outperform typical UEP schemes with or without channel interleaving. It is also observed that the performance of COSQ-based methods degrades gracefully (as is typical of such systems) in the presence of BER mismatch, while a powerful UEP method based on variable-length coding breaks down.

APPENDIX

In the following, we briefly explain the number of operations for rate allocation needed for our methods. In the IntraBlock method, $S = 13$ subbands are formed. Then, the variance of the subbands are computed and multiplied by the training distortion of a Laplacian source, trained for the given channel conditions. Finally, the rate is allocated in a bit-by-bit fashion, to the sub-band which maximizes the reduction in the overall distortion. In the InterBlock method, the image is first divided into $L \times L$ blocks ($L = 8$ in our case) and there are a total of K blocks. Each sub-source contains one coefficient from each block (at a given position). For both methods, the same rate allocation algorithm is used but with different input. The number of operations for each method is as follows.

The IntraBlock Method

- $RC+S$ multiplications to find the variances $\sigma_i^2, i=1,...,S$;
- $S \times r_{\max}$ multiplications to find $\sigma_i^2 \times d_L(r_i), i=1,...,S, r_i = 1,...,r_{\max}$;
- $S \times r_{\max}$ subtractions to find $\Delta D_L(r_i) \stackrel{\triangle}{=} \sigma_i^2 \left(d_L(r_i-1) - d_L(r_i)\right), i = 1,...,S, r_i = 1,...,r_{\max}$. $(d_L(0)=1)$;
- $S-1 \times RCr/(\frac{R}{16}\frac{C}{16}) = 256(S-1)r$ comparisons to find the best sub-source to allocate one bit (one with maximum $\Delta D_L(r_i), i=1,...,S$).

The InterBlock Method

- $RC+L^2$ multiplications to find the variances $\sigma_i^2, i=1,...,L^2$;
- $L^2 \times r_{\max}$ multiplications to find $\sigma_i^2 \times d_L(r_i), i=1,...,L^2, r_i = 1,...,r_{\max}$;
- $L^2 \times r_{\max}$ subtractions to find $\sigma_i^2 \left(d_L(r_i-1) - d_L(r_i)\right), i = 1,...,L^2, r_i = 1,...,r_{\max}$;
- $L^2-1 \times RCr/(\frac{R}{L}\frac{C}{L}) = L^2(L^2-1)r$ comparisons to find the best sub-source to allocate one bit.

Fig. 5. Decoded image of various IntraBlock methods at 0.5 bpp and BER = 0.1. ¿From top to bottom: The best tandem coding scheme (PSNR = 27.4 dB), COSQ-IL (PSNR = 23.5 dB), and the COSQ-based system designed for noisy channels with memory (PSNR = 28.6 dB).

References

[1] F. Alajaji and T. Fuja, "A communication channel modeled on contagion," *IEEE Trans. Inform. Theory*, vol. 40, pp. 2035-2041, Nov. 1994.

[2] F. Behnamfar, F. Alajaji, and T. Linder, "Progressive image communication over binary channels with additive bursty noise," in *Proc. Data Compression Conf.*, Snowbird, UT, Apr. 2002, pp. 272-281.

[3] V. Chande and N. Farvardin, "Progressive transmission of images over memoryless noisy channels," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 850-860, June 2000.

[4] Q. Chen and T. R. Fisher, "Image coding using robust quantization for noisy digital transmission," *IEEE Trans. Image Processing*, vol. 7, pp. 496-505, Apr. 1998.

[5] Q. Chen and K. P. Subbalakshmi, "Joint source-channel decoding for MPEG-4 video transmission over wireless channels," *IEEE J. Select. Areas Commun.*, vol. 21, pp. 1780-1789, Dec. 2003.

[6] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The JPEG2000 image coding system: an overview," *IEEE Trans. Consumer Electron.*, vol. 46, pp. 1103-1127, Nov. 2000.

[7] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels," *Bell Syst. Tech. J.*, vol. 42, pp. 1977-1997, Sept. 1963.

[8] N. Farvardin, "A study of vector quantization for noisy channels," *IEEE Trans. Inform. Theory*, vol. 36, pp. 799-809, July 1990.

[9] N. Farvardin and V. Vaishampayan, "On the performance and complexity of channel optimized vector quantizers," *IEEE Trans. Inform. Theory.*, vol. 37, pp. 155-160, Jan. 1991.

[10] S. Gadkari and K. Rose, "Robust vector quantizer design by noisy channel relaxation," *IEEE Trans. Commun.*, vol. 47, pp. 1113-1116, Aug. 1999.

[11] E. N. Gilbert, "Capacity of burst noise channels," *Bell Syst. Tech. J.*, vol. 39, pp. 1253-1265, Sept. 1960.

[12] R. Iordache, I. Tabus, and J. Astola, "Robust index assignment using Hadamard transform for vector quantization over finite-memory contagion channels," *Circuits, Systems and Signal Processing*, vol. 21, pp. 483-507, Sept.-Oct. 2002.

[13] R. Iordache, I. Tabus, and A. Beghdadi, "Media-based post-processing for VQ image transmission over noisy channels," in *Proc. Int. Symp. Signal Processing and its Applications*, Kuala Lumpur, Malaysia, Aug. 2001, pp. 48-51.

[14] J. D. Johnston, "A filter family designed for use in quadrature mirror filter banks," in *Proc. Int'l Conf. Acoustics, Speech, and Signal Processing*, Denver, 1980, pp. 291-294.

[15] J. Lim and D. L. Neuhoff, "Joint and tandem source-channel coding with complexity and delay constraints," *IEEE Trans. Commun.*, vol. 51, pp. 757-766, May 2003.

[16] D. Miller and K. Rose, "Combined source-channel vector quantization using deterministic annealing," *IEEE Trans. Inform. Theory*, vol. 42, pp. 347-356, Feb./Mar./Apr. 1994.

[17] M. Mushkin and J. Bar-David, “Capacity and coding for the Gilbert-Elliott channel,” *IEEE Trans. Inform. Theory*, vol. 35, pp. 1277-1290, Nov 1989.

[18] A. Nosratinia, J. Lu, and B. Aazhang, “Source-channel rate allocation for progressive transmission of images,” *IEEE Trans. Commun.*, vol. 51, pp. 186-196, Feb. 2003.

[19] N. Phamdo, F. Alajaji, and N. Farvardin, “Quantization of memoryless and Gauss-Markov sources over binary Markov channels,” *IEEE Trans. Commun.*, vol. 45, pp. 668-675, June 1997.

[20] C. Pimentel, T. H. Falk, and L. Lisbôa, “Finite-state Markov modeling of correlated Rician channels,” in *Proc. Canadian Workshop Inform. Theory*, Waterloo, Canada, May 2003, pp. 16-19.

[21] M. J. Ruf and J. W. Modestino, “Rate-distortion performance for joint source and channel coding of images,” in *Proc. Int’l Conf. Image Processing*, Washington D.C., Oct. 1995, vol. II, pp. 77-80.

[22] A. Said and W. A. Pearlman, “A new, fast and efficient image codec based on set partitioning in hierarchical trees,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, pp. 243-250, June 1996.

[23] P. G. Sherwood and K. Zeger, “Progressive image coding on noisy channels,” in *Proc. Data Compression Conf.*, Snowbird, UT, Mar. 1997, pp. 72-81.

[24] P. G. Sherwood, X. Tian, and K. Zeger, “Channel code block length and rate optimization for progressive image transmission,” in *Proc. Wireless Commun. Networking Conf.*, New Orleans, LA, Sept. 1999, pp. 978-982.

[25] K. P. Subbalakshmi and J. Vaisey, “Optimal decoding of entropy-coded Markov sources over channels with memory,” in *Proc. The 33rd Annual Conf. Inform. Sci. Syst.*, vol. 2, Mar. 1999, pp. 624-629.

[26] N. Tanabe and N. Farvardin, “Subband image coding using entropy-coded quantization over noisy channels,” *IEEE J. Select. Areas Commun.*, vol. 10, pp. 926-943, June 1992.

[27] N. Thomos, N. V. Boulgouris, M. G. Strintzis, “Wireless image transmission using turbo codes and optimal unequal error protection,” in *Proc. Int’l Conf. Image Processing*, Washington D.C., Sept. 2003, vol. I, pp. 73-76.

[28] V. Vaishampayan and N. Farvardin, “Optimal block cosine transform image coding for noisy channels,” *IEEE Trans. Commun.*, vol. 38, pp. 327-336, Mar. 1990.

[29] K. Zeger, J. Vaisey, and A. Gersho, “Globally optimal vector quantizer design by stochastic relaxation,” *IEEE Trans. Signal Processing*, vol. 40, pp. 310-322, Feb. 1992.